Effects of Uncertainties in Components on the Survival of Complex Systems with Given Dependencies

Axel Gandy
University of Hohenheim and University of Ulm
Germany
agandy@mathematik.uni-ulm.de

Abstract

When considering complex systems (i.e. systems composed of multiple components) in practice, the failure behaviour of the components is usually not known precisely. This is particularly true in early development phases of a product. We study the influence of uncertainties in the marginal distributions of the component's failure time on one-dimensional properties of the system's failure time (e.g. expectation, quantiles). We do not assume that the components are independent; instead we require that their dependence is given by a known copula. We consider two approaches:

In the first approach we assume that the margins have a bounded distance from known distributions. This approach leads to bounds on the one-dimensional properties and requires the solution of a non-trivial optimization problem. We provide solutions for some special cases.

The second approach is Bayesian and assumes some prior distribution on the marginal distributions. For example, one may assume that the margins belong to parametric classes and that distributions on the parameters are given. Simulation can be used to obtain the distribution of one-dimensional properties of the system's failure time.

1 Introduction

We consider a classical complex system allowing just two states: working (coded as 1) and failed (coded as 0). We assume that the system consists of n components and that the state of the components determines the state of the system, i.e. we have a structure function $\Phi : \{0,1\}^n \to \{0,1\}$. In this paper, we assume Φ to be monotone, meaning that Φ is monotone in each component, $\Phi(0,\dots,0) = 0$ and $\Phi(1,\dots,1) = 1$. The nonnegative random failure times of the components are denoted by $T_1,\dots,T_n$. Let $X_i(t) = 1_{\{T_i>t\}}$ be the state of the ith component at time t, where 1 denotes the indicator function. The probability that the system has failed up to a given time t is $F^S(t) := \mathrm{P}(\Phi(X_1(t),\dots,X_n(t)) = 0)$. For further discussion of complex systems we refer to (Aven and Jensen 1999).

If H denotes the joint cumulative distribution function of the failure times $T_1,\dots,T_n$ and $F_1,\dots,F_n$ denote their marginal distributions by Sklar's theorem there exists an n-copula C such that $H(t_1,\dots,t_n) = C(F_1(t_1),\dots,F_n(t_n))$. Recall that an n-copula is an n-variate cumulative distribution function with margins that are uniformly distributed on $[0,1]$. If $F_1,\dots,F_n$ are continuous then C is uniquely determined. If $T_1,\dots,T_n$ are independent then C can be chosen as the product copula $\Pi(x_1,\dots,x_n) = x_1x_2\dots x_n$. For further details we refer to (Nelsen 1999) and (Joe 1997).

In practice, H is usually not known precisely. This is particular true in early development phases. Often, one is interested in one-dimensional properties of the system distribution F^S like the expectation or quantiles. These properties are defined by mappings from the the space of cumulative distribution functions of nonnegative random variables, which we call D, into $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty,\infty\}$. Say $q : D \to \overline{\mathbb{R}}$ is one of these mappings. We want to study how imprecise knowledge of H influences $q(F^S)$. To simplify this task we restrict ourselves to uncertainties about the marginal distributions $F_1,\dots F_n$ and assume that the dependence structure of H is given by a known copula C. We consider two approaches. In the first approach we assume that we know $G_1,\dots,G_n \in D$ such that $d_i(F_i,G_i) \le \epsilon_i$ for some $\epsilon_i \ge 0$ where $d_1,\dots d_n$ are functions measuring distances in the space D. To derive the resulting bounds on $q(F^S)$ we have to solve a non-trivial optimization problem. We consider this approach in section 3.

The second approach is Bayesian in nature. We assume that the marginal distributions $F_1, \dots, F_n$ depend on some parameters $\theta_1, \dots, \theta_n$. These parameters do not have to be one-dimensional. We assume that the distribution of $(\theta_1, \dots, \theta_n)$ is known. In section 4 we consider this approach and give an example. In section 5 we will compare aspects of the two approaches.

Before starting with the two approaches we show in section 2 how the marginal distributions $F_1, \dots F_n$ can be separated from the structure function Φ and the copula C in the computation of F^S.

Proofs of the results can be found in (Gandy 2004).

2 System Reliability with dependent Components

We will introduce a function $G_{\Phi,C}$ and show how, together with the marginal distributions, it determines the cdf of the system F^S. Let $\Phi : \{0,1\}^n \to \{0,1\}$ be a monotone structure function and let C be an n-dimensional copula. Let $\tilde{C}$ be the probability measure on $([0,1]^n, \mathcal{B}([0,1]^n))$ induced by C. For $t \in [0,1]$ let $B_0^t := [0,t]$ and $B_1^t := (t,1]$. Let $G_{\Phi,C} : [0,1]^n \to [0,1]$ be given by

$$G_{\Phi,C}(t_1, \dots, t_n) := 1 - \sum_{\boldsymbol{x} \in \{0,1\}^n} \Phi(\boldsymbol{x}) \tilde{C}\left(\prod_{i=1}^n B_{x_i}^{t_i}\right).$$

Lemma 1. *Let $T_1, \dots, T_n$ be nonnegative random variables with marginal distributions $\boldsymbol{F} = (F_1, \dots, F_n) \in D^n$ and with copula C. For $t \in \mathbb{R}_+ = [0, \infty)$ let $X_i(t) := 1_{\{t < T_i\}}$. Let Φ be a structure function and $F^S \in D$ the cumulative distribution function of the complex system having structure Φ and components with lifetimes $T_1, \dots, T_n$, that is $F^S(t) = \mathrm{P}(\Phi(X_1(t), \dots, X_n(t)) = 0)$.*

Then it follows that $F^S(t) = G_{\Phi,C}(\boldsymbol{F}(t))$, for each t, i.e.

$$F^S = G_{\Phi,C} \circ \boldsymbol{F}.$$

The following lemma is a consequence of the assumption that Φ is monotone and of properties of copulas.

Lemma 2. *$G_{\Phi,C}$ is nondecreasing and continuous.*

Example 3. For a parallel system with n components, i.e. the system fails when all components have failed, $\Phi(x_1, \dots, x_n) = 1 - \prod_{i=1}^n (1 - x_i)$ and hence $G_{\Phi,C} = C$.

In the case of a serial system, i.e. the system fails when a single component fails, $\Phi(x_1, \dots, x_n) = \prod_{i=1}^n x_i$. If there are just $n = 2$ components then $G_{\Phi,C}(t_1, t_2) = t_1 + t_2 - C(t_1, t_2)$.

Figure 1 illustrates $G_{\Phi,C}$ for $n = 2$ components.

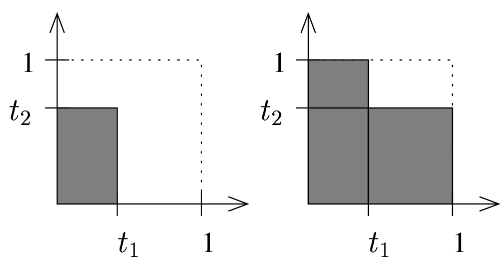


Figure 1: Illustration of $G_{\Phi,C}$ in the case $n = 2$. The copula C induces a measure on $[0,1]^2$. The value of $G_{\Phi,C}(t_1, t_2)$ is the measure of the shaded area in the case of a parallel/serial system (from left to right)

Example 4. If the product copula $\Pi(\boldsymbol{u}) = \prod_{i=1}^n u_i$ is used, i.e. the components are independent, then

$$G_{\Phi,\Pi}(t_1, \dots, t_n) = 1 - \sum_{\boldsymbol{x} \in \{0,1\}^n} \Phi(\boldsymbol{x}) \prod_{i=1}^n t_i^{1-x_i}(1 - t_i)^{x_i}$$

3 Bounds on the Margins

In this section, we want to study how one-dimensional properties of the system's cumulative distribution function behave if only bounds on the margins are known. More formally, let Φ be a monotone structure function, let C be an n-copula and suppose we know that the marginal distributions $F_1, \dots, F_n \in D$ of the failure times satisfy $d_i(F_i, G_i) \leq \epsilon_i, i = 1, \dots, n$, where $G_1, \dots, G_n \in D$ are known cumulative distribution functions, $\epsilon_i \geq 0$ and $d_i : D \times D \to \overline{\mathbb{R}}_+$ measure distances between cumulative distribution functions. Examples for d_i are the supremal distance $d_\infty(F, H) := \sup_{t \in \mathbb{R}_+} |F(t) - H(t)|$, L_p-distances on the cumulative distribution function $d_p(F, H) := \left(\int_0^\infty |F(t) - H(t)|^p dt\right)^{1/p}$ (for some $p > 0$), and L_p-distances on the inverse cumulative distribution function $d_p^{-1}(F, H) = \left(\int_{(0,1)} (F^{-1}(t) - H^{-1}(t))^p dt\right)^{1/p}$ (for some $p > 0$). d_2^{-1} is called Mallow's metric.

We are interested in a one-dimensional property of the system's cumulative distribution function which we assume to be given by a function $q : D \to \overline{\mathbb{R}}$. Examples for q include the expectation $E(F) := \int_0^\infty 1 - F(t) dt$ and quantiles $Q_p(F) := \inf\{t \in \mathbb{R}_+ : F(t) \geq p\}$ (for some $0 < p \leq 1$).

In practice it is interesting to know the minimal possible value of $q(F^S)$ given the restrictions on the margins. For this, the following optimization problem over $\boldsymbol{F} = (F_1, \dots, F_n)$ has to be solved.

$$(*) \begin{cases} q(F^S) = q(G_{\Phi,C} \circ \boldsymbol{F}) \to \min \\ \boldsymbol{F} \in D^n \\ d_i(F_i, G_i) \leq \epsilon_i, i = 1, \dots, n \end{cases}$$

We are not aware of a general solution of $(*)$. However, for some important special cases solutions can be given.

For the following we use the usual stochastic ordering on D, i.e. $F \leq G$ iff $F(x) > G(x)$ for all $x \in \mathbb{R}_+$.

Proposition 1. *If q is nondecreasing and $d = d_\infty$ then $\boldsymbol{F^0}$ given by*

$$F_i^0(t) := (G_i(t) + \epsilon_i) \wedge 1$$

is a solution of $(*)$.

This is a consequence of the monotonicity of Φ and q. Note that the functions E and Q_p are nondecreasing.

Next, we consider $(*)$ if $q = Q_p$, that is if we are interested in quantiles.

Proposition 2. *Suppose that for each i, the distance d_i has the property that if $H_0, H_1, H_2 \in D$ and $\forall t \in \mathbb{R}_+ : |H_0(t) - H_1(t)| \leq |H_0(t) - H_2(t)|$ then $d_i(H_0, H_1) \leq d_i(H_0, H_2)$.*
Furthermore, suppose that $q = Q_p$ for some $0 < p \leq 1$.
For $i = 1, \dots, n$, let $G_i^{s,z}(t) := G_i(t) \vee z 1_{[s,\infty)}(t)$, $z_i(s) := \sup\{z \in [0,1] : d_i(G_i, G_i^{s,z}) \leq \epsilon_i\}$ and $\boldsymbol{G}^s := (G_1^{s,z_1(s)}, \dots, G_n^{s,z_n(s)})$.

Then a lower bound for the optimal target value of $(*)$ is given by $t_0 := \inf\{t \in \mathbb{R}_+ : G_{\Phi,C}(\boldsymbol{G}^t(t)) \geq p\}$.

To actually comupte t_0, often one has to resort to numerical methods.

Solutions for the special cases of parallel and serial systems with independent components when $d_i(F, H) = \int_0^\infty |F(x) - H(x)|\, dx$ and $q = E$ are also possible.

4 Bayesian approach

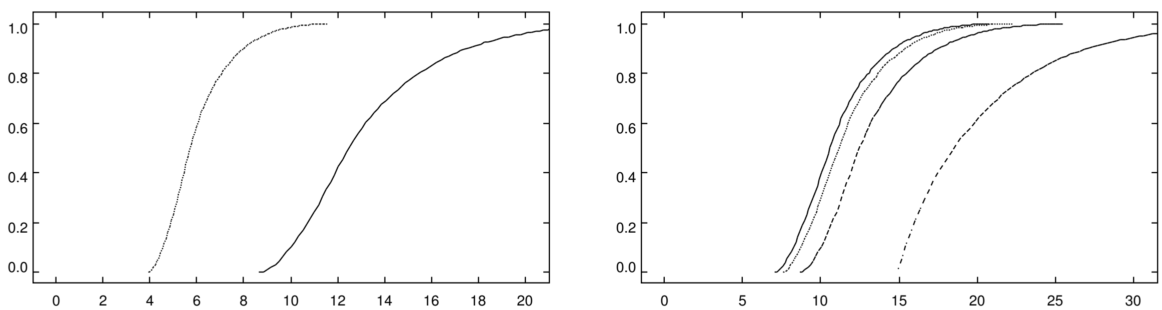
The second approach is Bayesian in nature. We assume that the marginal distributions $F_1, \dots F_n$ depend on some parameters $\theta_1, \dots, \theta_n$. These parameters do not have to be one-dimensional. We assume that the distribution of $(\theta_1, \dots, \theta_n)$ is known.

As a result of this approach one will get a distribution of $q(F^S)$. Explicit formulas for this distribution in the general case cannot be expected. However, using simulation one can get an impression of this distribution.

Example 5. Consider a 2-component serial system. Assume that the joint distribution of the failure times follows a Marshall-Olkin distribution, i.e. $T_1 = \min(Z_1, Z_{12})$, $T_2 = \min(Z_2, Z_{12})$, where Z_1, Z_2, Z_{12} are independent and exponentially distributed with rates $\lambda_1, \lambda_2, \lambda_{12}$. The copula of the joint distribution of T_1 and T_2 is a *generalized Cuadras-Augé* copula $C_{\alpha,\beta}(u,v) = \min(u^{1-\alpha}v, uv^{1-\beta})$ where $\alpha = \lambda_{12}/(\lambda_1 + \lambda_{12})$ and $\beta = \lambda_{12}/(\lambda_2 + \lambda_{12})$. Details of this can be found in (Nelsen 1999, Chapter 3.1.1).

Assume that we know that $C_{0.5,0.5}$ is the copula of our system and that T_1 and T_2 are exponentially distributed with rates θ_1 and θ_2, where θ_1 and θ_2 are i.i.d. following a uniform distribution on $[0.005, 0.015]$. Then the cumulative distribution function of the 0.1-quantile and the 0.2-quantile of the system is shown in the left diagram of Figure 2. If we replace $C_{0.5,0.5}$ by the Fréchet-Hoeffding lower bound copula $W(u,v) = \max(u+v-1, 0)$, the product copula $\Pi(u,v) = uv$ or the Fréchet-Hoeffding upper bound copula $M(u,v) = \min(u,v)$ the cumulative distribution function of the 0.2-quantile of the system is as shown in the right diagram of Figure 2.

Figure 2: Left: cdf of $Q_{0.1}(F^S)$ and $Q_{0.2}(F^S)$ of a 2-component serial system with copula $C_{0.5,0.5}$.
Right: cdfs of $Q_{0.2}(F^S)$ of the same system with copulas (from left to right) W, Π, $C_{0.5,0.5}$ and M.



5 Comparisons

A problem of the approach of section 3 is that it only yields a bound. This bound may be too pessimistic. Furthermore, in order to compute this bound one has to solve a non-trivial optimization problem. The advantage is that one does not only consider marginal distributions within a certain parametric class as in the approach of section 4. The advantage of the Bayes approach is that it yields a distribution of properties of the system's failure behaviour which may be more realistic than just a fixed bound. Since this distribution can be evaluated using simulation the Bayes approach is relatively straightforward to apply. Furthermore, the Bayes approach can be generalized to incorporate uncertainties in the copula.

Acknowledgements

Support of the "Deutsche Forschungsgemeinschaft" is gratefully acknowleded.

References

Aven, T. and U. Jensen (1999). *Stochastic models in reliability.* Springer, New York.

Gandy, A. (2004). Effects of uncertainties in components on the survival of complex systems with given dependencies. Technical report, University of Hohenheim. http://www.uni-hohenheim.de/~jensen/spp/publ.html.

Joe, H. (1997). *Multivariate Models and Dependence Concepts.* Chapman&Hall.

Nelsen, R. B. (1999). *An introduction to copulas*, Volume 139 of *Lecture Notes in Statistics.* Springer, New York.